二，三の確率分布に対する近似公式およびその応用

(抜取検査における損失函数の計算と通信呼理論における確率計算の簡易法)

真 壁 肇*

§1. 緒　　論

著者[5,9,10]および森村[9,10]は過去において，いくつかの確率分布を極限分布によっておきかえることの出来る近似公式を研究し，あわせてこのとき生ずる誤差の評価を行った．最近，著者はこの研究を更に進め[6,7]，いくつかの成果をえたが，ここではその中からとくに OR に関係あるものと考えられる抜取問題の損失函数の計算および通信呼理論への応用についてのべ，この近似公式を at hand に利用する二，三の方法を提案した．

§2. 用いられる近似公式

よく知られているように，2項分布

$$b(k\,;\,n,\,p)=\binom{n}{k}p^k(1-p)^{n-k}$$

は $np=\lambda=$一定 とし，$n\to\infty$, $(p\to 0)$ とすればポアソン分布

$$p(k\,;\,\lambda)=\frac{\lambda^k e^{-\lambda}}{k!}$$

で近似される．しかし，一般にこの近似は $np<5$ かつ $p<1/10$ でないと実用上は使用出来ないといわれている．これに対して，山内[13]，Nair[11] らの研究もあるが，著者ら[7),10] はポアソン分布の階差式を用いた補正を行うとよいことを見出し，その誤差の評価をも行った．すなわち，

[定理]

2項分布 $b(k\,;\,n,\,p)$ において $p<1/2$ とすれば

(1)
$$\sum_{k=0}^{c} b(k\,;\,n,\,p)=\sum_{k=0}^{c}\binom{n}{k}p^k(1-p)^{n-k}$$

$$=\sum_{k=0}^{c} p(k\,;\,\lambda)-\frac{\lambda^2}{2n}\triangle^2 P(c\,;\,\lambda) \qquad \text{(第1近似)}$$

$$-\frac{\lambda^3}{3n^2}\triangle^3 P(c\,;\,\lambda)+\frac{\lambda^4}{8n^2}\triangle^4 P(c\,;\,\lambda) \qquad \text{(第2近似)}$$

$$+R_2 \qquad \text{(誤 差 項)}$$

である．ここで

---

* 横浜国立大学　昭和 37 年4月 26 日受理

(2) $$\lambda = np$$

(3) $$\begin{cases} \triangle P(c;\lambda) = \sum_{k=0}^{c} p(k;\lambda) - \sum_{k=0}^{c-1} p(k;\lambda) = p(c;\lambda), \\ \triangle^2 P(c;\lambda) = \triangle P(c;\lambda) - \triangle P(c-1;\lambda), \\ \triangle^{n+1} P(c;\lambda) = \triangle^n P(c;\lambda) - \triangle^n P(c-1;\lambda) \end{cases}$$

であって，$p<1/10$ として計算すれば

(4) $$|R_2| < 62\,p^3,$$

とくに第1近似までとったときの誤差 R_1 は

(5) $$|R_1| < 5\,p^2$$

となり，補正項を全く用いないときの誤差 R は

(6) $$|R| < p$$

となる．

この誤差評価は二，三の観点より検討すると，この形としては best possible なものに近いことがわかっている[6,7]．しかしながら，あとでも述べるように実際に生ずる誤差と比較するとこれでもまだ相当大きいようである．したがって，この評価式は数学的興味はあるが，実用上の見地からはなお検討を要するものと思う．§3 でこれについて論ずることにする．

最後に，上の定理は負の2項分布に対してもそのまま適用できることを注意したい．すなわち

(7) $$\sum_{k=0}^{c} \pi(k;h,d) = \sum_{k=0}^{c} \frac{h(h+d)\cdots(h+\overline{k-1}\,d)}{k!}(1+d)^{-\frac{h}{d}-k}$$

$$= \sum_{k=0}^{c} p(k;h) + \frac{dh}{2}\triangle^2 P(c;h) \qquad \text{(第1近似)}$$

$$-\frac{d^2h}{3}\triangle^3 P(c;h) + \frac{d^2h^2}{8}\triangle^4 P(c;h) \qquad \text{(第2近似)}$$

$$+R_2 \qquad \text{(誤 差 項)}$$

となる．したがって，2項分布の場合と比べて

(8) $$p \leftrightarrow d, \quad \lambda \leftrightarrow h, \quad n \leftrightarrow h/d$$

とすれば，対応がついていることになる．この事実は著者が 1961 年に細い近似計算とともに発表したが，その後 1960 年の G. P. Patil による2項分布と負の2項分布との関連についての研究[2]によってもこのことが裏づけられることがわかった．

§3. 近似式の利用法とその近似度について

著者は前節にのべた近似公式を利用するのに便利なようにポアソン分布表からその階差表を作ることにした．用いたポアソン分布表は統計数値表[3]（第1表）のものである．つぎにその一例を示す．実際に計算作製したのは $m=0.1\sim1(0.1)$，$1\sim20(1)$ についてである．（ここでの m は λ のことである．）

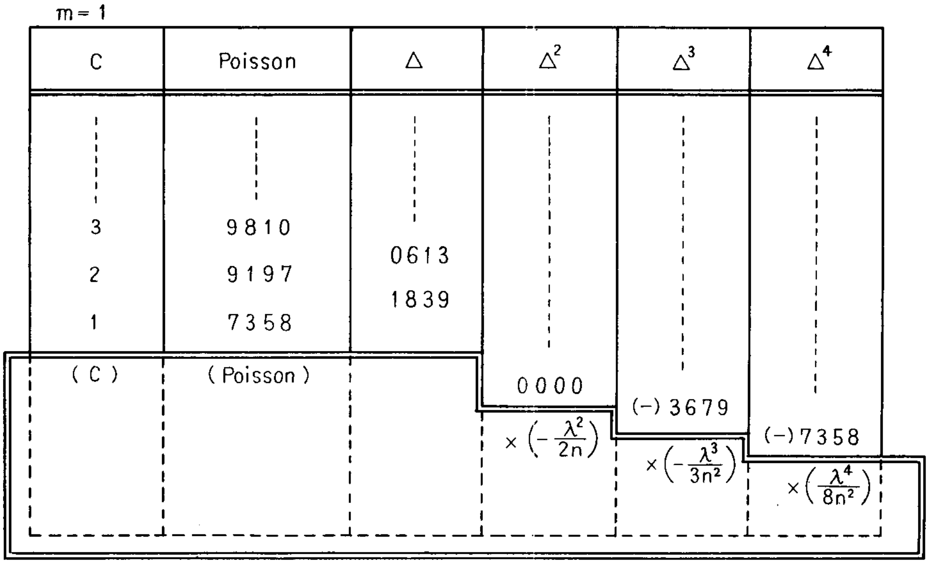
これだけの数表があれば第1図のように二重線の枠によって表わされた紙を用意し，それに

第 1 表

$m=1$

c	Poisson 値	△	△₂	△³	△⁴
6	9.999				
		0.005			
5	9.994		(−)0.026		
		0.031		0.096	
4	9.963		(−)0.122		(−)0.242
		0.153		0.338	
3	9.810		(−)0.460		(−)0.428
		0.613		0.761	
2	9.197		(−)1.226		0.152
		1.839		0.614	
1	7.358		(−)1.840		2.454
		3.679		(−)1.840	
0	3.679		0.000		1.839
		3.679		(−)3.679	
−1	0.000		3.679		(−)7.358
		0.000		3.679	
−2	0.000		0.000		3.679
		0.000		0.000	
−3	0.000		0.000		0.000

$\times\left(-\dfrac{\lambda^2}{2n}\right)$ の値を書き込んでおけば $\sum_{k=0}^{c} b(k\,;\,n,\,p)$ の計算に役立つことがわかる.



第 1 図

著者は実際にこのようにして近似計算を実行した. この結果

1) 第1近似を用いたときには

$$n=4\sim40,\quad p<\text{約 }15\%$$

2) 第2近似を用いたときには

$$n=10\sim50,\quad p<\text{約 }30\%$$

または

$$n=5\sim10,\quad p<25\%$$

でも実用上はまず差支えないとの結論をえた．ここで実用上差支えないとの基準は近似度が小数点以下4桁目が完全に一致するとは期待はできないとしても，3桁目にやっと響きうる程度という意味である．一例を示そう(第2表)．$m=\lambda=np=1$，$n=4$，$p=0.25$ の場合を考える．これは上にのべた最悪の条件に対応する近似度のチェックと見てよい．

第2表

c	Poisson 値	第 1 近 似	第 2 近 似	2項分布和
4	9963	1	1	1
3	9810	9963	9957	9961
2	9197	9427	9493	9492
1	7358	7358	7378	7383
0	3679	3223	3174	3164

(Poisson 値は統計数値表[3]による)

最終桁は丸め誤差の影響で上の表程度以上改善することは出来ない．

負の2項分布の場合にも同様のことが期待できる．上にのべた事実を

$$p \to d, \quad n \to h/d \qquad (\lambda=h)$$

として適用すればよい．

なお著者の計算した誤差評価の式(4)～(6)と比べると事実上の誤差はこれらの式による値の約1/8～1/50 程度である．

§4. 通信呼理論における応用例

小島[4]によれば，即時式完全線群を取扱うに当って

入線数；N，　出線数；n，　呼量；a

とすれば，r ケの呼の存在する確率 $p(r)$ は

(9)
$$p(r)=\binom{N}{r}\alpha^r \Big/ \sum_{k=0}^{n}\binom{N}{k}\alpha^k, \quad \alpha=\frac{a}{N-a}$$

で与えられるとされている．同氏はこれを Martin の式とよび，さらにその数値計算に言及している．

(9)において

$$分子=\binom{N}{r}\left(\frac{\alpha}{\alpha+1}\right)^r\left(\frac{1}{\alpha+1}\right)^{N-r}\cdot(\alpha+1)^N$$

$$分母=\sum_{k=0}^{n}\binom{N}{k}\left(\frac{\alpha}{\alpha+1}\right)^k\left(\frac{1}{\alpha+1}\right)^{N-k}\cdot(\alpha+1)^N$$

とすれば

(10)
$$p(r)=\binom{N}{r}\left(\frac{\alpha}{\alpha+1}\right)^r\left(\frac{1}{\alpha+1}\right)^{N-r} \Big/ \sum_{k=0}^{n}\binom{N}{k}\left(\frac{\alpha}{\alpha+1}\right)^{N-k}$$

$$\fallingdotseq \{p(r\,;\,N\alpha/\overline{\alpha+1})+\triangle_1\}\Big/\{\textstyle\sum_{k=0}^{n}p(k\,;\,N\alpha/\overline{\alpha+1})+\triangle_2\}$$

となる。ここで

$$\text{(11)}\qquad \begin{cases} \triangle_1=-\dfrac{(N\alpha/\overline{\alpha+1})^2}{2N}\triangle^2 p(r\,;\,N\alpha/\overline{\alpha+1}) \\[2ex] \triangle_2=-\dfrac{(N\alpha/\overline{\alpha+1})^2}{2N}\triangle^2 P(n\,;\,N\alpha/\overline{\alpha+1}) \end{cases}$$

とおいた.

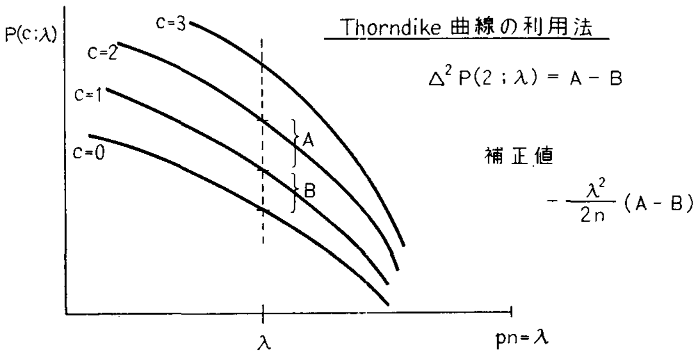
一例として，小島[4)]のとりあげた case;

$$N=20,\qquad n=10,\qquad a=5$$

を考えると

$$\alpha/\overline{\alpha+1}=0.25,\qquad N\alpha/\overline{\alpha+1}=5$$

となる．つまり，$n=20$, $p=1/4$, $\lambda=np=5$ の2項分布に対する近似と考えて計算すれば第2図のようになることがわかる．この事実は第1近似式を用いると $p(r)$ の値は略完全に真の値と一致していることを教えている.



第 2 図

著者のしらべたところによれば

$$\alpha/\overline{\alpha+1}<0.4$$

でなるならば，第2近似式を用いれば $p(r)$ の計算の目的は充分に達せられるようである.

その他この公式を適用しうるものとしては

たとえば，呼損率

$$B=\frac{\binom{N-1}{n}\alpha^n}{\sum_{r=0}^{n}\binom{N-1}{r}\alpha^r}$$

とか r の分散

$$\sigma_r{}^2=a_c(1-a_c)+\frac{N(N-1)\alpha^2\sum_{r=0}^{n-1}\binom{N-2}{r}\alpha^r}{\sum_{r=0}^{n}\binom{N}{r}\alpha^r},$$

a_c; 平均同時接続数[4)]

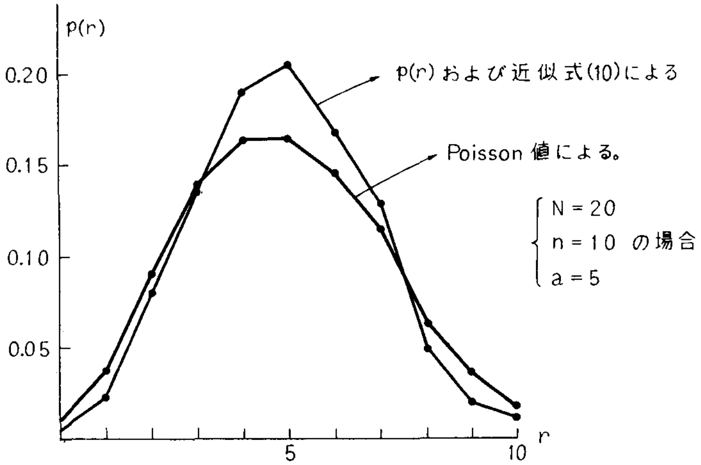
の計算への応用等が考えられる.

[注意]

1. B の計算に当って, B の値が小さいとき, この近似公式では相対誤差の保証が充分ではないので, この点はまだチェックを必要とするように思われる.

2. 第1近似公式を用いるには, Thorndike 曲線(たとえば日本科学技術連盟; 抜取検査教程; 附, 数表 p. 210)を利用すると便利である.

つまり, 第3図のようにするとよい.



第 3 図

§ 5. 受入抜取検査における損失函数の計算

大きさ N のロットを受入検査によって購入しているメーカーを考えよう. このメーカーはもしロットを受入れるならば, そのロットの中に含まれる不良品によって, 1個当り a 円の損失をうけるし, またロットを不合格させれば, 全数検査による選別を行うので, 1個当り b 円の cost を必要とする.

したがって, 1個の抜取検査費用を1円として基準化しておけば, 損失函数は $N\gg n$ として,

(12) $$K=n+a\cdot N\times(\text{出検品質})+b\cdot N\times(\text{不合格の確率})$$

となる. ここで n は抜取個数, 出検品質はロットの不良率を p, そのロットの合格する確率を $L(p)$ とすれば $pL(p)$ となる.

ところで Hamaker[2)] はこの計算ではロットの不良率 $p5$ が一定であるという不自然な点がともなうとして,

$$K=n+a\cdot N\int_0^1 pL(p)\cdot\varphi(p)dp+b\cdot N\int_0^1(1-L(p))\varphi(p)dp \tag{13}$$

$\varphi(p)$; ロット不良率の分布

を用いるべきであることを提案している．著者の研究したところによれば，$\varphi(p)$は一般にガンマ分布に従うことがわかっているが[8]，この分布は平均(ロット平均不良率) $\bar{p}$ と変動係数(ロット不良率の変動係数)v_p を $v_p=1, \sqrt{2}/2, 1/2, 1/3$ としておさえれば，その形をほぼつくすことが出来る．著者[8]はこの K を $n\bar{p}$=一定，$\bar{p}$; 小 として計算したところ

$$K=n+aN\bar{p}\sum_{k=0}^{c}\frac{(h+d)(h+2d)\cdots(h+kd)}{k!}(1+d)^{-\frac{h+d}{h}-k} \tag{14}$$

$$+bN\sum_{k=c+1}^{\infty}\frac{h(h+d)\cdots(h+\overline{k-1}d)}{k!}(1+d)^{-\frac{d}{h}-k}$$

ここで　$\bar{p}$; ロット平均不良率

$h=n\bar{p}$,

$\sqrt{d/h}=\sqrt{v_p}$

c; 合格判定個数

をえた．Hald[1] も同じ形の式を出しているが，まだ，この式を利用して抜取方式を求めるには至っていない．

いま，$\sqrt{v_p}=1/2$ とすれば，$h/d=4$ となるから $h\leqq 1$, $d\leqq 1/4$ なるとき(14)の各項は近似式(7)で有効に評価出来る．一般に $\sqrt{v_p}\leqq 1/2$ のときは(7)式が有用であることは§1 の誤差の検討によって明かなところである．また $\sqrt{v_p}>1/2$ のときは $\sqrt{v_p}=1, \sqrt{v_p}=\sqrt{2}/2$ として計算すると(14)は割合簡単な式(例えば(15))で表わせることがわかる．ここではとくに一例として $\sqrt{v_p}=1$ のときの，最適解の求め方についてのべることにする．

いま $\sqrt{v_p}=1$ とすれば $h=d$ であるから(14)にこれを代入すれば

$$K=n+aN\bar{p}\sum_{k=0}^{c}(k+1)(1+h)^{-2-k}+bN\sum_{k=c+1}^{\infty}(1+h)^{-1-k} \tag{15}$$

$$=n+aN\bar{p}\left\{1-(c+2)\left(\frac{h}{1+h}\right)^{c+1}+(c+1)\left(\frac{h}{1+h}\right)^{c+2}\right\}+bN\left(\frac{h}{1+h}\right)^{c+1}$$

となるが，ここで $h=n\bar{p}$ として

$$\frac{\partial}{\partial n}K=0$$

をとけば，各 c の値を定めたときの $K=\min$ となる n が定まる．これから c の値を動かして $K_{\min}$ の最小値を読めば (n, c) が最適解としてえられることになる．(ただし K の第 2, 3 項の近似式は実際は僅かであるが $O(\bar{p})$ なる誤差をもっていてそれが N 倍されるから，求まった値に対しては，吟味を最後に加える必要がある)

［例］　1.　$N=1000$,　$\bar{p}=\frac{1}{100}$,　$a=50$,　$b=1/2$　とすれば，

最適解　　$n=67$,　$c=0$

2.　$N=5000$,　$\bar{p}=\frac{1}{100}$,　$a=50$,　$b=1$　のとき,

最適解　　$n=118$,　$c=3$　　　　となる.

§6. む　す　び

以上では，2項分布および負の2項分布がポアソン分布で近似されることが OR においてはどのような計算の簡易化をもたらすかを考えてみた.

著者は上の近似理論を更に拡張して，ベルヌイ試行列において各試行時の事象出現確率が異る場合にも(ここで k 回目の試行においてはその確率を p_k とする)，$\max p_k=\alpha$ が小さくかつ $\sum_{k=1}^{n} p_k=\lambda$ が一定として(1)と同じような式を導いてある[7]. このような結果は，OR のいろいろな問題に利用されるのではないかと思われるので(例えば統計学のサンプリング問題にはこの応用例もあるが)何かよい例を御教示いただければ望外の幸である.

最後に，御指導を賜った朝香鉄一先生，国沢清典先生，木暮正夫先生および畏友森村英典氏に厚く御礼申し上げます. また，数値計算についていろいろお世話いただいた反町洋一氏および山崎善光氏に御礼申し上げます.

参　考　文　献

(1) HALD, A.; The Compound Hypergeometric Distribution and a System of Single Sampling Inspection Plans Based on a Prior Distributions and Costs, Technometrics **2** (1960)

(2) HAMAKER, H. C.; Some Basic Principles of Sampling Inspection by Attributes, Applied Statistics **7** (1958)

(3) 北川敏男，増山元三郎；新編統計数値表，河出書房(1952)

(4) 小島哲；通信呼理論の研究，科学新興社(1949)

(5) MAKABE, H; A Normal Approximation to Binomial Distribution, Rep. Stat. App. Res., JUSE **4** (1955)

(6) MAKABE, H; On the Approximation to Some Limiting Distributions with Applications, to appear in Kōdai Math. Sem. Rep.

(7) MAKABE, H.; On the Approximation to some Limiting Distributions with applications to the Theory of the Sampling Inspection by Attributes, to appear in Kōdai Math. Sem. Rep.

(8) 真壁肇；計数抜取検査における抜取方式の改良に関する研究，東京工業大学(1961)

(9) MAKABE, H and MORIMURA, H; On the Approximation to some Limiting Distributions, Kōdai Math. Sem. Rep. **8** (1956)

(10) MAKABE, H. and MORIMURA, H.; A Normal Approximation to Poisson Distribution, Rep. Stat. App. Res., JUSE **4** (1955)

(11) NAIR, K. R.; The Studentized Form of the Extreme Mean Square Test in the Analysis of Variance, Biometrika **35** (1948)

(12) PATIL, G. P.; On the Evaluation of the Negative Binomial Distribution with Example, Technometrics **2** (1960)

(13) 山内二郎；不完全ベータ函数の近似式について，(1952)